Diabetologia (2009) 52:844–849
DOI 10.1007/s00125-009-1281-3

ARTICLE

Investigation of DNA polymorphisms in *SMAD* genes for genetic predisposition to diabetic nephropathy in patients with type 1 diabetes mellitus

A. J. McKnight · A. M. Woodman · M. Parkkonen · C. C. Patterson · D. A. Savage · C. Forsblom · K. A. Pettigrew · D. Sadlier · P. H. Groop · A. P. Maxwell · The Warren 3/UK GoKinD Study Group

Received: 1 December 2008 / Accepted: 13 January 2009 / Published online: 27 February 2009

Abstract
Aims/hypothesis SMAD proteins are involved in multiple signalling pathways and are key modulators of gene expression. We hypothesised that genetic variation in selected *SMAD* genes contributes to susceptibility to diabetic nephropathy.
Methods We selected 13 haplotype tag (ht) single nucleotide polymorphisms (SNPs) from 67 variants identified by resequencing the *SMAD2* and *SMAD3* genes. For *SMAD1, SMAD4* and *SMAD5* genes, genotype data were downloaded for 217 SNPs from Phase II of the International HapMap project. Of these, 85 SNPs met our inclusion criteria, resulting in the selection of 13 tag SNPs for further investigation. A case–control approach was employed, using 267 nephropathic patients and 442 controls with type 1 diabetes from Ireland. Two further populations (totalling 1,407 patients, 2,238 controls) were genotyped to validate initial findings. Genotyping was conducted using iPLEX, TaqMan and gel electrophoresis.
Results The distribution of genotypes was in Hardy–Weinberg equilibrium. Analysis by the χ^2 test of genotype and allele frequencies in patients versus controls in the Irish population (n=709) revealed evidence for the association of one allele at 5% level of significance (rs10515478, $p_{\text{uncorrected}}$=0.006; $p_{\text{corrected}}$=0.04). This finding represents a relatively small difference in allele frequency of 6.4% in the patient group compared with 10.7% in the control group; this difference was not supported in subsequent investigations using DNA from European individuals with similar phenotypic characteristics.
Conclusions/interpretation We selected an appropriate subset of variants for the investigation of common genetic risk factors and assessed *SMAD1* to *SMAD5* genes for association with diabetic nephropathy. We conclude that common polymorphisms in these genes do not strongly influence genetic susceptibility to diabetic nephropathy in white individuals with type 1 diabetes mellitus.

Keywords Association · Genetic predisposition · Polymorphism · SMAD · Susceptibility gene

Electronic supplementary material The online version of this article (doi:10.1007/s00125-009-1281-3) contains supplementary material, which is available to authorised users.

A. J. McKnight (✉) · A. M. Woodman · D. A. Savage · K. A. Pettigrew · A. P. Maxwell
Nephrology Research Group, Queen's University of Belfast, c/o Regional Genetics Centre, Level A, Tower Block, Belfast City Hospital, Lisburn Road, Belfast BT9 7AB, Northern Ireland, UK
e-mail: a.j.mcknight@qub.ac.uk

M. Parkkonen · C. Forsblom · P. H. Groop
Folkhälsan Institute of Genetics, Biomedicum Helsinki, Helsinki, Finland

M. Parkkonen · C. Forsblom · P. H. Groop
Division of Nephrology, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland

C. C. Patterson
Epidemiology Research Group, Queen's University of Belfast, Belfast, UK

D. Sadlier
Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland

Abbreviations

ACR	Albumin/creatinine ratio
BMP	Bone morphogenic protein
ESRD	End-stage renal disease
FinnDiane	Finnish Diabetic Nephropathy Study
GoKinD	Genetics of Kidneys in Diabetes
ht	Haplotype tag
HWE	Hardy–Weinberg equilibrium
LD	Linkage disequilibrium
MAF	Minor allele frequency
r-SMAD	Receptor SMAD
SNP	Single nucleotide polymorphism
UAER	Urinary albumin excretion rate

Introduction

Diabetic nephropathy is the commonest cause of end-stage renal disease (ESRD), contributing to increased morbidity and mortality rates in patients with diabetes. Despite extensive efforts, the genetic risk factors contributing to the complex aetiology of diabetic nephropathy have not been identified [1].

SMAD proteins are integral components of TGFβ and bone morphogenic protein (BMP) signal transduction pathways. SMAD1 and SMAD5 are BMP-regulated receptor SMADs (r-SMADs) with SMAD4 classified as a common SMAD [2]. SMAD2 and SMAD3 are TGFβ/activin-regulated r-SMADs that promote epithelial to mesenchymal transitions and fibrogenesis. Stimulation triggers the phosphorylation and translocation of r-SMADs, thus facilitating interaction with SMAD4 to transcriptionally modulate gene expression [3]. Elevated SMAD1 levels are associated with early renal dysfunction and SMAD1 levels in glomeruli are correlated with progression of renal disease [4]. In a mouse model of diabetic nephropathy, inhibition of the TGFβ–SMAD2/3 pathway resulted in reduced renal fibrosis [5]. In a small family-based study, two single nucleotide polymorphisms (SNPs) in SMAD3 were associated with diabetic nephropathy [6]. SMAD4 plays a key role in the SMAD cascade for TGFβ and BMP signalling, and SMAD5 is involved in vasculogenesis and angiogenesis [2]. SMADs have critical functions and we hypothesised that genetic variation in genes encoding SMAD proteins described above would promote or inhibit mechanisms by which TGFβ and BMP contribute to diabetic nephropathy.

This report describes the comprehensive evaluation of common polymorphisms in the genomic regions encompassing five biological candidate genes (*SMAD1*, *SMAD2*, *SMAD3*, *SMAD4*, *SMAD5*) for association with diabetic nephropathy.

Methods

Participants Ethical approval was obtained from the appropriate Research Ethics Committees and all participants gave written informed consent before the study was conducted. Population heterogeneity was minimised by using comparable phenotypes in all three populations genotyped (Table 1). The initial population used genomic DNA samples from 709 white individuals with parents and grandparents born in Ireland. Nephropathy patients (n=267) demonstrated persistent proteinuria (>0.5 g/24 h) developing more than 10 years after initial diagnosis of diabetes, and also had hypertension (>135/85 mmHg and/or treatment with antihypertensive agents) and retinopathy. Controls without nephropathy (n=442) demonstrated no evidence of renal disease after at least 15 years of type 1 diabetes mellitus and were not being prescribed antihypertensive medication. Type 1 diabetes was diagnosed on the basis of insulin dependence from diagnosis and an age of onset of less than 35 years.

The second, replicate population comprised 1,099 white individuals with type 1 diabetes from UK and includes individuals recruited as part of the UK Warren 3/Genetics of Kidneys in Diabetes (GoKinD) collections for diabetic nephropathy [7]. Recruitment criteria were identical to those used in the Irish study, with the exceptions that type 1 diabetes was diagnosed at age of onset of less than 31 years and that participants were third-generation born in UK. All patients with type 1 diabetes recruited as controls had consistently normal urinary albumin excretion rates (UAERs). Of the patients with diabetic nephropathy, 27% were receiving renal replacement therapy (dialysis or renal transplant) and due to their ESRD status did not have UAER formally assessed. For the remaining nephropathy patients proteinuria was assayed either by urinary albumin/creatinine ratio (ACR) or by 24 h urine protein excretion measurements. Where UAER was determined by ACR, the mean value was 118.4 mg/mmol. All nephropathy patients had a minimum ACR of >30 mg/mmol. For the remaining nephropathy patients where UAER was measured by 24 h urine protein excretion, the mean value was 5.7 g/24 h and all nephropathy patients had a minimum urine protein excretion of >0.5 g/24 h.

The third population comprised 2,565 white individuals with type 1 diabetes from the nationwide Finnish Diabetic Nephropathy Study (FinnDiane) [8]. In the FinnDiane cohort, there were 1,736 control individuals with normal UAER (<30 mg/24 h or <20 μg/min) and 810 nephropathy patients, of whom 543 were macroalbuminuric (UAER >300 mg/24 h or 200 μg/min in at least two of three consecutive timed urine collections) and 267 had progressed to ESRD. The diagnosis of type 1 diabetes in FinnDiane was based on age at onset of diabetes <35 years

Table 1 Clinical characteristics of genotyped populations

Characteristic	All Ireland collection		UK Warren 3/GoKinD collection		FinnDiane collection		
	Nephropathy patients	Control	Nephropathy patients	Control	Nephropathy patients	Control[a]	Control[b]
n	267	442	597	502	810	1,070	666
Age at diagnosis (years)	16.7±11.5	15.0±8.1	14.5±7.6	15.6±8.0	11.6±7.0	14.2±8.2	19.0±8.1
BMI (kg/m^2)	25.7±4.8	26.3±4.2	26.4±4.8	26.1±4.2	25.3±3.9	25.0±3.2	24.6±3.3
Duration of diabetes (years)	32.0±9.3	27.2±9.3	33.9±9.3	29.2±9.1	30.2±8.2	27.2±9.2	9.4±3.3
HbA_{1c} (%)	10.6±1.5	9.5±1.2	8.6±1.7	8.2±1.4	8.9±1.6	8.2±1.3	8.4±1.5
Systolic BP (mmHg)	147±20	124±14	136±30	127±16	147±21	132±17	126±13
Diastolic BP (mmHg)	84±11	76±7	81±12	75±9	84±11	78±9	78±9

Values are mean±SD

[a] Individuals with duration of diabetes ≥15 years

[b] Individuals with duration of diabetes <15 years

and insulin treatment initiated within 1 year of diagnosis. Antihypertensive medication was prescribed for 93.5% of the nephropathy patients, for 20.7% of control participants with type 1 diabetes duration >15 years and for 5.4% of control participants with diabetes duration <15 years.

SNP selection To account for functional genetic elements adjacent to all genes of interest, genomic regions of interest were extended to 3 kb upstream from the reference mRNA transcription start site and 3 kb downstream of the stop codon. Overlapping PCR fragments (~500 bp) were designed to screen *SMAD2* and *SMAD3* genes for variants. Genes of 46 individuals (23 nephropathy patients; 23 controls) were resequenced on an genetic analyser (ABI 3730; Applied Biosystems, Foster City, CA, USA) to obtain allele frequencies for the population. This provided 95% power to detect all polymorphisms with greater than 5% minor allele frequency (MAF). Pairwise linkage disequilibrium (LD) was evaluated using D′ and haplotype frequencies estimated using SNPHAP and haplotype tag (ht) SNPs identified by htsearch command in STATA (www-gene.cimr.cam.ac.uk/clayton/software/stata, accessed 8 August 2008). With the release of Phase II of the International HapMap Project [9], selection of tag SNPs became a cost-effective option. Sample-based genotypes (data release number 20) were downloaded for all variants in genetic regions surrounding *SMAD1*, *SMAD4* and *SMAD5*. As the study groups under investigation are white, downloaded genotypes were restricted to those for the CEPH (Utah residents with ancestry from northern and western Europe; CEU) population. Haploview [10] was used to visualise LD and haplotype blocks between these variants. Tag SNPs were selected using a pairwise tagging approach where r^2 exceeded 0.8 for all downloaded SNPs with minimum MAF >5% and no deviation from Hardy–Weinberg equilibrium (HWE) was seen at $p<0.001$ [11].

Genotyping Genotyping was performed for all variants in Irish and UK samples using iPLEX (Sequenom, San Diego, CA, USA), TaqMan (Applied Biosystems), agarose gel electrophoresis or direct capillary sequencing. Nephropathy patients and controls were randomly arranged in 384-well format with four negative controls, twelve duplicate samples per plate and four father–mother–proband trios used for experimental quality control. Unique dbSNP identifiers are provided for all variants (Table 2; Electronic supplementary material [ESM] Table 1) and full details of primers and reaction conditions are available from the authors. Genotyping of the provisionally associated SNP (rs10515478) was confirmed in the Irish population and investigated in Finnish samples by a commercially available TaqMan assay (c_11256545_10).

Statistical analysis Genotype frequencies were assessed for HWE using a χ^2 goodness-of-fit test. Genotype and allele frequencies were compared using the χ^2 test with the level of statistical significance set at 5%. Haplotype analysis was conducted where SNPs formed a haplotype block, using 95% CIs on D' [12]. Permutation (n=100,000) adjustment for multiple testing was performed in Haploview [10].

Results

We identified 24 variants in *SMAD2* (chr18q21.1: NC_000018.8_ 43613464..43711510), of which 21 were SNPs and three deletions. Twelve variants were novel and two htSNPs were identified. Screening of *SMAD3* (chr15q22.33: NC_000015.8_ 65145249..65274587) identified 43 variants, of which 12 were novel; 11 htSNPs were required to evaluate *SMAD3*.

Of the 217 SNPs retrieved from HapMap, 85 met our inclusion criteria, with tag SNP analysis identifying 13 markers

Table 2 Genotype and allele counts for all variants genotyped in the Irish population

SNP/ Genotype	Nephropathy patients (*n*)	Controls (*n*)	Allele	Nephropathy patients (*n*)	Controls (*n*)
*SMAD1*_rs1016792					
TT	184	286	T	441	713
TC	73	141	C	93	171
CC	10	15			
*SMAD1*_rs12505085					
AA	152	258	A	401	670
AG	97	154	G	131	214
GG	17	30			
*SMAD1*_rs13120931					
GG	64	128	G	269	457
GT	141	201	T	251	407
TT	55	103			
*SMAD1*_rs2118438					
GG	181	294	G	439	717
GA	77	129	A	89	159
AA	6	15			
*SMAD1*_rs6537355					
AA	202	337	A	463	769
AG	59	95	G	65	107
GG	3	6			
*SMAD1*_rs714195					
GG	88	161	G	309	512
GA	133	190	A	215	354
AA	41	82			
*SMAD2*_rs2000709					
CC	80	130	C	286	457
CT	126	197	T	244	413
TT	59	108			
*SMAD2*_ rs34998689					
++	210	362	+	471	799
+–	51	75	–	55	75
– –	2	0			
*SMAD3*_ rs58048858					
AA	114	180	A	333	561
AC	105	201	C	199	319
CC	47	59			
*SMAD3*_ rs62005980					
GG	185	317	G	437	750
GA	67	116	A	79	128
AA	6	6			
*SMAD3*_rs11635753					
TT	162	263	T	413	687
CT	89	161	C	117	197
CC	14	18			

Table 2 (continued)

SNP/ Genotype	Nephropathy patients (*n*)	Controls (*n*)	Allele	Nephropathy patients (*n*)	Controls (*n*)
*SMAD3*_ss107911942					
– –	226	399	–	484	834
–+	32	36	+	36	38
++	2	1			
*SMAD3*_rs2289261					
CC	117	213	C	363	602
CG	129	176	G	157	270
GG	14	47			
*SMAD3*_rs35874463					
AA	225	369	A	491	805
AG	41	67	G	43	79
GG	1	6			
*SMAD3*_rs2289791					
GG	137	240	G	375	654
GT	101	174	T	145	210
TT	22	18			
*SMAD3*_rs11556089					
GG	232	358	G	494	784
AG	30	68	A	36	78
AA	3	5			
*SMAD3*_rs11556090					
GG	73	134	G	278	489
GA	132	221	A	250	389
AA	59	84			
*SMAD3*_rs12900401					
CC	225	363	C	484	795
CT	34	69	T	40	75
TT	3	3			
*SMAD3*_rs1052488					
TT	137	205	T	367	584
TC	93	174	C	147	270
CC	27	48			
*SMAD4*_rs10502913					
GG	138	251	G	385	662
GA	109	160	A	141	206
AA	16	23			
*SMAD4*_rs8084630					
GG	97	158	G	323	535
GA	129	219	A	207	347
AA	39	64			
*SMAD4*_rs948588					
GG	230	374	G	494	811
GA	34	63	A	36	67
AA	1	2			

Table 2 (continued)

SNP/ Genotype	Nephropathy patients (*n*)	Controls (*n*)	Allele	Nephropathy patients (*n*)	Controls (*n*)
*SMAD5*_rs10515478[a]					
CC	233	355	C	498	789
CG	32	79	G	34	95
GG	1	8			
*SMAD5*_rs7707640					
TT	130	198	T	367	588
TC	107	192	C	157	292
CC	25	50			
*SMAD5*_rs746994					
CC	170	270	C	418	683
CT	78	143	T	295	193
TT	12	25			
*SMAD5*_rs17749249					
TT	186	293	T	441	707
TA	69	121	A	79	153
AA	5	16			

[a] p=0.006, OR 0.57 (95% CI 0.37–0.87)

that would effectively evaluate common sequence variation in the genetic sequence surrounding and encompassing *SMAD1* (chr4q31: NC_000004.10 146622401..147060850), *SMAD4* (chr18q21.1: NC_000018.8 46810611..46860145) and *SMAD5* (Chr5q31: NC_000005.8 135496435..135594458) (ESM Fig. 1). The distribution of all genotypes was in HWE with >99% genotyping accuracy and >97% of genotypes successfully called. Analysis by the χ^2 test of genotype and allele frequencies between nephropathy patients and controls in the Irish collection revealed evidence for association of one variant at the 5% level of significance (rs10515478; nominal allelic p=0.006; p_{adj}=0.04; OR 0.57; 95% CI: 0.37–0.87). Further genotyping of this SNP in a UK population revealed a difference in allele frequencies of only 1.2% between nephropathy patients and controls (p=0.5). Similarly, MAF in the Finnish population differed by less than 2% between nephropathy patients (13.6%) and diabetic controls (11.7%). No significant differences were observed for any other SNP investigated. Haplotype analyses (ESM Fig. 2) did not reveal a statistically significant association with diabetic nephropathy.

Discussion

We used proxy SNPs (based on correlations between causal mutations and haplotypes) to investigate association with diabetic nephropathy by exploiting inherent SNP redundancy to maximise genotyping efficiency and minimise costs. This indirect approach necessitated <10% of annotated SNPs being typed for these genes, with a high degree of allelic capture due to the strong LD observed in these *SMAD* genes. Provisional evidence for association was observed for a single marker (rs10515478) in the Irish population. Original genotypes for this SNP were confirmed in all samples using an alternative technology to minimise systematic genotyping error. With the recent availability of DNA from the national, multi-centre Warren 3/GoKinD UK collection for diabetic nephropathy, rs10515478 was efficiently genotyped in this larger population.

As the initial association observed for rs10515478 with diabetic nephropathy was not confirmed in the second population, a third DNA collection was genotyped to further elucidate the relationship. The provisional association failed to replicate in multiple study populations with more than 90% power to detect an association, in individual populations, that would be equivalent to that of the magnitude observed in the original study group. The pooled replication collection had more than 85% power to detect a risk allele (p<0.05), with an OR of 1.25, assuming a MAF of 10% in controls.

The lack of replication may be due to different patterns of LD (and hence different tag SNPs) across different populations [13]; however, this is unlikely as HapMap tag SNPs have been shown to be transferable across European ancestry [14]. The lack of independent replication may be due to epistasis [13], but it is most likely that the original finding was a false positive association highlighted by chance as a result of a lower MAF in a smaller sample size. The sample size of the initial Irish collection provided 90% power to detect an OR equal to 1.75 with a MAF of 10%. To minimise the risk of a false negative finding in our population, we subsequently genotyped all SNPs in the recently available UK collection, increasing power to more than 90% to identify a risk allele with OR of 1.5. Contrary to a previous study reporting weak association for two SNPs in *SMAD3* from 72 European trios (p>0.03<0.05) [6] and in agreement with a recent finding [15], no significant differences were observed between nephropathy patients and controls in our larger study population.

Coverage of genomic regions comprised an average of one SNP per 0.8 kb, with common variation in HapMap assessed with a mean r^2 of 0.92, 0.995 and 0.99 for *SMAD1*, *SMAD4* and *SMAD5* respectively. htSNPs were identified with a minimum r^2 of 0.84 for *SMAD2* and 0.90 for *SMAD3*. It is possible that rare alleles (MAF<5%) or allelic heterogeneity within these genes are true disease risk factors for diabetic nephropathy, but many thousands of individuals would be required to detect such a significant association or small effect and such phenotypically similar collections of DNA are presently limited. It is also possible that genetic variants in more distant regulatory regions of these genes may influence the development of diabetic nephropathy.

Determining definitive genetic risk factors for diabetic nephropathy has proved challenging. However, the availability of larger sample collections and collaborative projects, as well as the increasing cost-effectiveness of higher-throughput technologies, are making it easier to draw more definitive conclusions. Using an efficient approach, we comprehensively assessed the role of common genetic variation in five candidate genes for association with diabetic nephropathy. The large population tested does not support the hypothesis that common genetic variation in *SMAD1*, *SMAD2*, *SMAD3*, *SMAD4* or *SMAD5* strongly influences genetic susceptibility to diabetic nephropathy in white individuals with type 1 diabetes mellitus.

Acknowledgements This work was financially supported by the Northern Ireland Kidney Research Fund and the Northern Ireland Research and Development Office. Genotyping of UK Warren 3/GoKinD DNA samples was performed by S. Heggarty of Queen's University Belfast Core Technology Unit. The Warren 3/UK GoKinD Study Group was jointly funded by Diabetes UK and the Juvenile Diabetes Research Foundation. We are greatly indebted to all participants as well as the doctors and nurses who collected data for the FinnDiane Study (see ESM for list). The FinnDiane Study was supported by grants from the Folkhälsan Research Foundation, the Wilhelm and Else Stockmann Foundation, the Sigrid Juselius Foundation, the Signe and Arne Gyllenberg Foundation and the European Commission (EURAGEDIC QLG2-CT-2001-01669, ADDNET LSHB-CT-2003-503364 and DIANA LSHB-CT-2006-037681).

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

References

1. Savage DA, Bain SC, McKnight AJ, Maxwell AP (2007) Gene discovery in diabetic nephropathy. Curr Diab Rep 7:139–145
2. Attisano L, Tuen Lee-Hoeflich S (2001) The Smads. Genome Biol 2:REVIEWS3010
3. Chang H, Brown CW, Matzuk MM (2002) Genetic analysis of the mammalian transforming growth factor-beta superfamily. Endocr Rev 23:787–823
4. Abe H, Matsubara T, Iehara N et al (2004) Type IV collagen is transcriptionally regulated by Smad1 under advanced glycation end product (AGE) stimulation. J Biol Chem 279: 14201–14206
5. Petersen M, Thorikay M, Deckers M et al (2008) Oral administration of GW788388, an inhibitor of TGF-β type I and II receptor kinases, decreases renal fibrosis. Kidney Int 73:705–715
6. Ewens KG, George RA, Sharma K, Ziyadeh FN, Spielman RS (2005) Assessment of 115 candidate genes for diabetic nephropathy by transmission/disequilibrium test. Diabetes 54: 3305–3318
7. Currie D, Maxwell AP, Sadlier D, McKnight AJ, and The Warren 3/UK GoKinD Study Group (2008) Investigation of *Adducin 2* (beta) DNA polymorphisms in genetic predisposition to diabetic nephropathy in type 1 diabetes. Diabet Med 25:1001–1005
8. Mäkinen V-P, Forsblom C, Thorn L et al (2008) Metabolic phenotypes, vascular complications and premature deaths in a population of 4,197 patients withtype 1 diabetes. Diabetes 57:2480–2487
9. Frazer KA, Ballinger DG, Cox DR et al (2007) A second generation human haplotype map of over 3.1 million SNPs. Nature 449:851–861
10. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265
11. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D (2005) Efficiency and power in genetic association studies. Nat Genet 37:1217–1223
12. Gabriel SB, Schaffner SF, Nguyen H et al (2002) The structure of haplotype blocks in the human genome. Science 296:2225–2229
13. Ioannidis JP, Boffetta P, Little J et al (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 37:120–132
14. Montpetit A, Nelis M, Laflamme P et al (2006) An evaluation of the performance of tag SNPs derived from HapMap in a Caucasian population. PLoS Genet 2:e27
15. Trégouet DA, Groop PH, McGinn S et al (2008) G/T substitution in intron 1 of the UNC13B gene is associated with increased risk of nephropathy in patients with type 1 diabetes. Diabetes 57:2843–2850